AD-A265 270

FASTC-ID(RS)T-0338-93

FOREIGN AEROSPACE SCIENCE AND TECHNOLOGY CENTER

CODING OBJECTS WITH THE USE OF A DIVIDING HYPERPLANE FOR THEIR CLASSIFICATION

by

Yu. V. Devingtal'

(Perm')

PARTIALLY EDITED MACHINE TRANSLATION

FASTC-ID(RS)T-0338-93 21 May 1993

MICROFICHE NR:

CODING OBJECTS WITH THE USE OF A DIVIDING HYPERPLANE FOR THEIR CLASSIFICATION

By: Yu. V. Devingtal'
(Perm')

English pages: 15

Source: Izvestiya Akademii Nauk SSSR Tekhnicheskaya Kibernetika, Nr. 3, 1971; pp. 139-147

Country of origin: USSR
This document is a machine translation.
Input by: Young H. Perry
Merged by: Young H. Perry
Requester: WL/MLIM/Dr. Al Jackson
Approved for public release; Distribution unlimited.

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

THIS TRANSLATION IS A RENDITION OF THE ORIGINAL FOREIGN TEXT WITHOUT ANY ANALYTICAL OR EDITORIAL COMMENT STATEMENTS OR THEORIES ADVOCATED OR IMPLIED ARE THOSE OF THE SOURCE AND DO NOT NECESSARILY REFLECT THE POSITION OR OPINION OF THE FOREIGN AEROSPACE SCIENCE AND TECHNOLOGY CENTER.	PREPARED BY: TRANSLATION DIVISION FOREIGN AEROSPACE SCIENCE AND TECHNOLOGY CENTER WPAFB, OHIO

U. S. BOARD ON GEOGRAPHIC NAMES TRANSLITERATION SYSTEM

Block	Italic	Transliteration	Block	Italic	Transliteration
А а	*А а*	A, a	Р р	*Р р*	R, r
Б б	*Б б*	B, b	С с	*С с*	S, s
В в	*В в*	V, v	Т т	*Т т*	T, t
Г г	*Г г*	G, g	У у	*У у*	U, u
Д д	*Д д*	D, d	Ф ф	*Ф ф*	F, f
Е е	*Е е*	Ye, ye; E, e*	Х х	*Х х*	Kh, kh
Ж ж	*Ж ж*	Zh, zh	Ц ц	*Ц ц*	Ts, ts
З з	*З з*	Z, z	Ч ч	*Ч ч*	Ch, ch
И и	*И и*	I, i	Ш ш	*Ш ш*	Sh, sh
Й й	*Й й*	Y, y	Щ щ	*Щ щ*	Shch, shch
К к	*К к*	K, k	Ъ ъ	*Ъ ъ*	"
Л л	*Л л*	L, l	Ы ы	*Ы ы*	Y, y
М м	*М м*	M, m	Ь ь	*Ь ь*	'
Н н	*Н н*	N, n	Э э	*Э э*	E, e
О о	*О о*	O, o	Ю ю	*Ю ю*	Yu, yu
П п	*П п*	P, p	Я я	*Я я*	Ya, ya

*ye initially, after vowels, and after ъ, ь; e elsewhere.
When written as ё in Russian, transliterate as yё or ё.

RUSSIAN AND ENGLISH TRIGONOMETRIC FUNCTIONS

Russian	English	Russian	English	Russian	English
sin	sin	sh	sinh	arc sh	$\sinh^{-1}$
cos	cos	ch	cosh	arc ch	$\cosh^{-1}$
tg	tan	th	tanh	arc th	$\tanh^{-1}$
ctg	cot	cth	coth	arc cth	$\coth^{-1}$
sec	sec	sch	sech	arc sch	sech^{-1}
cosec	csc	csch	csch	arc csch	csch^{-1}

Russian	English
rot	curl
lg	log

GRAPHICS DISCLAIMER

All figures, graphics, tables, equations, etc. merged into this translation were extracted from the best quality copy available.

CODING OBJECTS WITH THE USE OF A DIVIDING HYPERPLANE FOR THEIR CLASSIFICATION

Yu. V. Devingtal'

(Perm')

Methods of object coding are examined, if for their classification a dividing hyperplane is used. Methods of expanding the space of the codes are proposed, if in the initial space distribution by hyperplane is not possible. The obtained code tables, which were obtained, may be used for simple classification of objects without the use of computers. The coding methods indicated were used for the classification of geophysical objects and for the prediction of the properties of two-component alloys.

1. Preliminary observations. Let us consider the division of the set of objects into two classes, if each object of this set is described by n parameters, whose possible values are discrete. Virtually any parameter has discrete values, since even with continuous change in a certain value its values are quantized, for example, by the method of its measurement. Let the parameter with the number i have k_i values ($i=1, 2, ..., n$). Usually each value of the parameter is coded with a fixed form and the object is assigned by a set of codes of the values of its parameters. The decisive rule for the classification is constructed in the space of the codes. As was noted in study [1], an unsuccessful coding method can considerably complicate the classification problem. This article examines such coding method of objects, with which it is possible to accomplish the classification with the help of the simplest decisive rule - a hyperplane, moreover beforehand let us fix this hyperplane in the space of the codes. For simplification of the calculations during classification it is expedient to select a hyperplane, passing through the origin of the coordinates in the space of the codes and which has equal coefficients with all coordinates, i.e., the hyperplane of form

$$x_1+x_2+...+x_n=0. \tag{1.1}$$

Let us select the codes of the values of the parameters so that this hyperplane would completely divide the training sequence. In study [1] the solution method of this problem

is proposed, but is not very suitable, if the number of values of the parameters is sufficiently large. Furthermore, in the method the circumstance are not considered that in some problems the classifications of the property of objects continuously are changed together with a change in the parameters, i.e., the codes of these parameters must also be changed sufficiently smoothly.

We will assume, that each object from the training sequence is preset in the form of a set of n integers (numbers of the parameter values, which correspond to this object). Let the objects with numbers $r=1, 2, ..., m_1$ belong to the 1st class, and with numbers $r=m_1+1, m_1+2, ..., m_1+m_2$ - to the 2nd class, thus, the training sequence is m_1+m_2 sets of n integers:

$$j_{1r}, j_{2r}, ..., j_{nr} \quad (r=1, 2, ..., m_1+m_2). \tag{1.2}$$

Page 140

2. Formulation of the problem. Let i be the number of the parameter ($i=1, 2, ..., n$), and j_i - the ordinal number of its values ($j_i=1, 2, ..., k_1$). We will search for such a table of codes $\{x_{ij}\}$ of the parameter values so that the codes would be the polynomials of a given degree p_i ($i=1, 2, ..., n$) from a number of the value. Then

$$x_{ij_i}=x_{i0}+\sum_{k=1}^{p_i} C_{j_i}^{k}\Delta_i^k x , \tag{2.1}$$

where x_{io} - is a certain value for forming the i-th column of the code table (value of the code at the zero point, i.e., when $j_i=0$), and $\Delta_i^k x$ - is the finite difference in the k-th order at the zero point for the i-th column of the table, $C_{j_i}^k$ - binomial coefficients, moreover $C_{j_i}^k$ = 0 when $k>j_i$.

Let us require that hyperplane (1.1) completely divide training sequence (1.2), i.e., it is necessary to select numbers $x_{ij_{ir}}$ so that the system of the inequalities would be satisfied:

$$\sum_{i=1}^{n} x_{ij_{ir}}<0 \qquad for\ r=1,2,\ldots,m_1$$

$$\sum_{i=1}^{n} x_{ij_{ir}}>0 \qquad for\ r=m_1,\ldots,m_1+m_2. \tag{2.2}$$

Let us name these two groups of inequalities 1st and 2nd class inequalities respectively. Taking into account condition (2.1) the system of inequalities (2.2) can be rewritten in the form

$$y_0+\sum_{i=1}^{n}\sum_{k=1}^{p_i} C_{j_{ir}}^{k}\Delta_i^k x<0(>0) \tag{2.3}$$

for

$$r=1,\ 2,\ \ldots,\ m_1,\ (r=m_1+1,\ \ldots,\ m_1+m_2),$$

where

$$y_0=\sum_{i=1}^{n} x_{i0}.$$

Thus, problem is of determination of $N+1$ of numbers y_0 and $\Delta_i^k x\left(N=\sum_{i=1}^{n} p_i\right)$,

which satisfy the system of inequalities (2.3). When $p_i=1$ (i=1, 2, ..., n) system (2.3) takes the form

$$y_0+\sum_{i=1}^{n} j_{i_r}\Delta_i x<0(>0),$$

for $r=1,\ 2,\ \ldots,\ m_1\ (r=m_1+1,\ \ldots,\ m_1+m_2)$,
which coincides with the type of the system of inequalities, which are obtained during the solution of the problem about the determination of coefficients of the dividing plane, if the codes of the parameter values coincide with their numbers. In this case the coefficients of the dividing plane coincide with numbers $\Delta_i x$ (i=1, 2, ..., n), and absolute term - with y_0.

When $p_i=k_i$, system (2.3) coincides with the system, proposed in [1]. In this case each

parameter value will receive its own code (the number of unknowns N will coincide with the number of parameter values). The degree p_i of the polynomial should be selected so that system (2.3) would be solved. As was shown in [1], even when $p_i = k_i$ the system (2.3) does not always have a solution. In this case it is useful to widen the number of parameters, which describe the object, counting for the new parameters of combination on 2, 3, ..., m of the values of initial. These new parameters also have the finite number of values.

Page 141

Let us name m with a depth of combinations. It is obvious that the number of unknowns in the code tables { x_{ij_i} } is very great during the utilization of combinations of initial parameters as new parameters. With depth m the number of parameters

$$n+\sum_{k=1}^{m} n_k=\sum_{k=1}^{m} C_n^k,$$

where $n_k = C_n^k$ - the number of combinations of the n parameters with respect to k, and the number of values, which can accept these parameters

$$N_m=\sum_{i=1}^{n} k_i+\sum_{i_2>i_1=1}^{n} k_{i_1}k_{1_2}+\ldots+\sum_{i_m>i_m-1>\ldots>i_1=1}^{n} k_{i_1}k_{i_2}\ldots k_{i_m}.$$

With $m=1$ we will obtain a set of initial parameters with the number of unknowns

$$N_1=\sum_{i=1}^{n} k_i .$$

Let us note that due to the enormous number of unknown values in the code tables it is expedient that during the use of combinations of the parameters of the certain depth m to no longer examine the combination of these parameters of a smaller depth. In this case the number of new parameters will be equal to n_m, and the number of elements in the code tables

$$N_m = \sum_{i_m > i_m - 1 > \ldots > i_1 = 1}^{n} k_{i_1} k_{i_2} \ldots k_{i_m} .$$

Subsequently we will consider this observation as carried out. In the expanded code space of the parameter values as the dividing hyperplane it is possible to take a hyperplane of form (1.1), where $n=n_m$. Moreover the codes of combinations z_{ij_r} ($r=1, 2, \ldots, m_1, \ldots, m_1+m_2$) must satisfy the system (2.2), where $n=n_m$.

The table of codes { x_{ij_i} } in this case will contain with large m many more values than are represented in system (2.2). Therefore the value of this table as a classification instrument will be reduced with an increase in m. At the same time there is such a value for m, with which system (2.2) will be resolved, if it was not resolved with $m=1$. Actually, with $m=n$ system (2.2) is always solved, if only one and the same object is not registered in different classes in the training sequence. In this case it is possible to consider the combination of all initial parameters a single parameter, which determines the object, and in the code table, which consists of one column, the lines, which only correspond to the objects of the training sequence are filled. It is possible to assume $x_{ij_r}=-1$ if $r=1, 2, \ldots, m_1$ and $x_{ij_i} = 1$ for $r=m_1+1, \ldots, m_1+m_2$. According to this table it will be possible to divide only the training sequence and it will not represent a value for the classification. Therefore the number of terms in the code table must be matched with the number of inequalities in system (2.2). During the use of combinations of the parameters as the new parameters it is possible also to superimpose on their codes of condition (2.1), in this case the system of inequalities (2.2) will take the form of (2.3).

Page 142

The combinations of the parameters were used in the program type "Kora" [kora - crust] [2] with $m=3$. The program "Kora" is seeking a certain solution of the system (2.2), namely, those of values $x_{ij_{ir}}$, which are encountered only in inequalities of one of the classes (and besides frequently enough), assume it is equal to -1 for the 1st (class), +1 for the 2nd

class and 0 for inequalities of both classes, which are encountered simultaneously. According to this code table objects of the 1st class with substitution of the code values into the left side of the equation of hyperplane (1.1) will give a negative result, and the 2nd (class) - positive. This method is used conveniently in the case of well separated classes, where some combinations of the values of three parameters sufficiently frequently belong only to one class. One method of solving the system of inequalities (2.2) with the use of combinations of the parameters for badly separated classes (i.e. when program "Kora" it gives bad results) is described in [3].

Thus, for the solution of the distribution problem by training in sequence (1.2) hyperplane (1.1) is necessary to solve the system of inequalities (2.2), and, if the values of the initial parameters are coded, it is expedient to superimpose further limitations on the smoothness of the values of the codes and to use system (2.3). With the insolvability of system (2.3) even when $p_i=k_i$ $(i=1, 2, ..., n)$ it is possible to use combinations of parameter values and to seek the codes of these combinations, solving system (2.2). Let us assume the values of the unknowns, which were not included in this system, to be equal to zero. For any training sequence, which does not have identical objects in different classes, it is possible to find such space, in which this training sequence is divided by a hyperplane of form (1.1). If we introduce positive numbers s_r $(r=1, 2, ..., m_1+m_2)$, system (2.2) can be written in the form

$$\delta_r \sum_{i=1}^{n} x_{ij_{ir}} = s_r , \tag{2.4}$$

where

$$\delta_r = \begin{cases} -1 \; for \; r=1,2,...,m_1, \\ \;\; 1 \; for \; r=m_1+1,...,m_1+m_2; \end{cases}$$

and system (2.3) in the form

$$\delta_r y_0+\delta_r\sum_{i=1}^{n}\sum_{k=1}^{p_i} C_{j_{ir}}^{k}\Delta_i^k x=s_r(r=1,2,\ldots,m_1+m_2). \tag{2.5}$$

3. Solution of system inequality. Let us assume that values s_i are known, then system (2.5) contains

$$N+1=\sum_{i=1}^{n} p_i+1$$

the unknowns: y_0, $\Delta_i^k x$ ($k=1, 2, \ldots, p_i$ $i=1,2, \ldots, n$).

Let us designate for the brevity these unknowns through $y_0, y_1, \ldots, y_n$, coefficients $\delta_r, \delta_r C_{j_{ir}}^{k}$, which correspond to unknowns y_i ($i=0, 1, \ldots, N$), through a_{ir} ($i=0, 1, \ldots, N$; $r=1, 2, \ldots, m_1+m_2$), moreover $a_{or}=\delta_r$. It is convenient to take the following numbering of unknowns:

$$y_0=\sum_{i=1}^{n} x_{i0}, \; y_1=\Delta_1 x,\; y_2=\Delta_2 x,\; \ldots,\; y_n=\Delta_n x,$$

$$y_{n+1}=\Delta_1^2 x,\; \ldots,\; y_{2n}=\Delta_n^2 x,\; \ldots,\; y_{pn}=\Delta_n^p x, \ldots .$$

Page 143

In this case

$$a_{ir}=\begin{cases}\delta_r & \text{with } i=0,\\ 0 & \text{with } [\frac{i-1}{n}] > \min(p_i j_{ir}),\ i>0,\\ \delta_r C_{j_{ir}}^{[\frac{i-1}{n}]+1} & \text{with } [\frac{i-1}{n}] \le \min(p_i j_{ir}),\ i>0,\end{cases}$$

$$N=p^{+}n, \text{ where } p^{*}=\max p_i.$$

The vector, whose components compose the coefficients of the line of system (2.5), let us designate through $A_r = \{a_{0r}, a_{1r}, ..., a_{nr}\}$, and let $Y = \{y_0, y_1, ..., y_n\}$. With such designations system (2.5) will take the form

$$(A_r, Y) = s_r, \tag{3.1}$$
$$s_r > 0 \ (r = 1, 2, ... m_1 + m_2).$$

The system (2.4) will take the same form, if we designate the unknown values of the codes in y_i, and the vector A_r will have n components, equal to one, and the rest - zero. If we designate $\min_r s_r = y_{N+1}$, then the solution of system (3.1) may lead to the solution of the problem of linear programming: to find the maximum of linear form $z = y_{N+1}$ within limitations $(A_r, Y) \geq y_{N+1}$ $(r=1, 2, ..., m_1+m_2)$. This problem corresponds to finding the Chebyshev point of system (2.2) with condition (2.1), which can be found with any method of solution of the linear programming problem [4]. In view of the uniformity of conditions one of the unknowns, for example y_0, can be fixed. The distribution problem will be solved, if max $z > 0$. The solvability of this problem will be explained in resolving the problem of linear programming. However, it is possible to propose another solution procedure.

Let us select from the vector system A_r $(y=1, 2, ..., m_1+m_2)$ the maximum linearly independent system. Let it consist of q vectors $q \leq m_1+m_2$. If $q = m_1+m_2$, then the solution vector Y can be represented in the form of a linear combination of vectors A_r,

$$Y = \sum_{r=1}^{m_1+m_2} \alpha_r A_r,$$

and to determine coefficients α_r from the system of equations

$$\sum_{r=1}^{m_1+m_2} \alpha_r (A_i, A_r) = s_r \quad (r=1,2,...m_1+m_2),$$

with arbitrarily selected right sides, for example when $s_r = 1$. The determinant of this system is Gram's determinant, which is positive with linear independence of vectors A_r. If vectors A_r in the training sequence are all linearly independent, then the problem of distribution is always solved.

Let us consider now case of $q < m_1+m_2$. Let the first q of the vectors A_r be linearly

independent. In this case the solution also may be sought in the form of a linear combination

$$Y=\sum_{i=1}^{q} \alpha_i A_i .$$

(3.2)

Moreover coefficients α_i will be solutions of system

$$\sum_{i=1}^{q} \alpha_i (A_i,A_r)=s_r \quad (r=1,2...,q).$$

(3.3)

Page 144

This system is solved with any s_r, therefore it is necessary to ensure this selection of numbers s_r, so that Y, obtained according to formula (3.2) from the solution of system (3.3), would satisfy and to the remaining equations of system (3.1), when $s_r>0$. Then it is possible to write system(3.1) in the form

$$\sum_{i=1}^{q} \alpha_i (A_i,A_r)=s_r \quad with\ r=1,2,...,q,$$

(3.4)

$$\sum_{i=1}^{q} \alpha_i (A_i,A_r)-s_r=0 \quad with\ r=q+1,...,m_1+m_2,$$

where s_r with $r=1, 2, ..., q$ we will consider it as given, but s_r with $r=q+1, ..., m_1+m_2$ - unknowns.

Let us solve system (3.4) with regard to unknowns s_r. The determinant of this system is different from zero and is equal to $(-1)^{m_1+m_2-q}\Delta_q$, where Δ_q - Gram's determinant of the first q vectors (they are, on our assumption, linearly independent). According to Cramer formulas we will obtain

$$s_{q+t}=\frac{1}{\Delta_q}\sum_{i=1}^{q}\Delta_{i,q+t}s_i \quad (t=1,2,\ldots,m_1+m_2-q),$$

where $\Delta_{i,q+t}$ - determinant of order q, which differs from Δ_q only in terms of the i-th line, in which are elements of form (A_i, A_1), (A_i, A_2)...(A_i, A_q) are substituted with elements (A_{q+t}, A_1) (A_{q+t}, A_2), ..., (A_{q+t}, A_q) i.e.

$$\Delta_{i,q+t}=\begin{vmatrix} (A_1,A_1) & (A_1,A_2) & \cdots & (A_1,A_q) \\ \cdots & \cdots & \cdots & \cdots \\ (A_{i-1},A_1) & (A_{i-1},A_2) & \cdots & (A_{i-1},A_q) \\ (A_{q+t},A_1) & (A_{q+t},A_2) & \cdots & (A_{q+t},A_q) \\ (A_{i+1},A_1) & (A_{i+1},A_2) & \cdots & (A_{i+1},A_q) \\ \cdots & \cdots & \cdots & \cdots \\ (A_q,A_1) & (A_q,A_2) & \cdots & (A_q,A_q) \end{vmatrix}$$

Let us designate $\Delta_{i,q+t}/\Delta_q=\beta_{i,\,q+t}$. Then

$$s_{q+t}=\sum_{i=1}^{q}\beta_{i,q+t}s_i \quad (t=1,2,\ldots,m_1+m_2). \qquad (3.5)$$

If system (3.5) has a solution, all components which are positive, then the distribution problem is solved. For the solution of the distribution problem it suffices to solve system (3.3), after substituting into the right side the appropriate components of the positive solution of system (3.5). But if system (3.5) does not have the positive solution, then the distribution problem is not solved.

Let us note some obviously sufficient signs of the existence (and of nonexistence) of a positive solution for system (3.5). System (3.5) has a positive solution, if at least one of the conditions are fulfilled:

1) $\beta_{i,\,q+t}\geq 0$ for all possible values of i and t,
2) will be located such $i=i_0$, for which $\beta_{i0,\,q+t}>0$ for all $t=1, 2, \ldots, m_1+m_2-q$.

System (3.5) does not have the positive solution, if is located at least one such $t=t_0$, for which $\beta_{i,q+t_0} \leq 0$ for i=1, 2, ..., q.

If system (3.5) does not satisfy any of the enumerated sufficient conditions, then let us use the following method: assume $s_i=1$ for $i=1, 2, ..., q$, if it proves that

$$s_{q+t}=\sum_{i=1}^{q} \beta_{i,q+t}>0 \quad with\ t=1,2,...,m_1+m_2-q,$$

that the problem is solved and system (3.5) has a positive solution.

Page 145

But if among the numbers s_{q+t}, in this case, there proves to be a negative, then let us select the line with smallest value $\sum_{i=1}^{q} \beta_{i,q+t}$ and solve it relative to s_j, coefficient $\beta_{j,\ q+t}$ of which is positive (this coefficient it will be located, it will otherwise fulfill a sufficient condition of the insolvability of problem) and let us substitute the obtained expression for s_j into the remaining equations. We check after this, is there no line with non-negative coefficients (then the problem is not permitted), otherwise we again repeat conversion described above. After a certain number of steps we will obtain a positive solution of system (3.5) or we will be convinced of the insolvability of the distribution problem with this method of coding. With the inseparability it is possible to either leave the training sequence points, which prevent distribution, or to increase a quantity of unknowns in system (2.2), after increasing the degree of the polynomials p_i or after widening a quantity of parameters, after introducing combinations.

After obtaining the solution of system (2.2), we will have a code table, suitable for recognition of objects, not included in the training sequence.

4. Example. Let us consider the construction of a table for the distribution of frozen and thawed rocks according to the data of geophysical measurements. Three parameters were measured in holes every 5 m: apparent resistance ρ_k, spontaneous polarization (PS) and a caliper log. Required to determine the propagation border of frozen rocks, i.e., to divide all layers into frozen (class I) and thawed (class II). The training sequence was

formed as follows: since the hole passed to the region of permafrost, then the first several layers can be considered the representatives of frozen layers, and with sufficient hole depth it is possible to assume that the lowest layers are thawed. According to the small training sequence, which was obtained, (4-5 representatives of each class) a classing table is constructed and is used for the distribution of the middle layers of hole. Usually, in this case several layers, adjacent to the upper, are confidently classified as frozen, and part of the lower - as thawed. The training sequence is widened due to these layers and the dividing table, which was already defined, is constructed again. After several steps we obtain a table, on which the border of the permafrost in projected sufficiently clearly, which is well coordinated with the experimental data. Let us consider how the tables for recognition were constructed. First of all let us index the values of the parameters. Range of change ρ_k - from 1 to 300 Ω. Therefore as the first value let us take 1 Ω and from and the remaining values let us index by steps of 1 Ω. For PS the border of change for this region - from -20 to 30; therefore the reference point (1st value) is equal to -20 (step 1). For the caliper log the 1st value of -25 (step 1). Then each object is a set of three integers, for example 15, 22, 18.

The degrees of the polynomials, with the help of which the code tables were constructed, were chosen experimentally so that on the available material recognition would be produced with the highest conformity; with the experimental data on some holes, accepted as a reference. It was determined, that it is expedient to take $p_1=3$, $p_2=2$, $p_3=1$. Thus, for construction of the divided tables it was necessary to find 7 values: the total initial value y_0, three differences - for the construction of the codes of impedance, two - for PS and one difference - for the caliper log. It is impossible to completely give the initial data table.

Page 146

Let us give only the fragment of this table

class I	class II
15, 17, 19	14, 24, 13
30, 26, 17	10, 25, 13
45, 25, 17	11, 33, 15
....	....

The corresponding vectors A_r will appear (for $p_1=3$, $p_2=2$, $p_3=1$) as follows:

-1, -15, -17, -19, -105, -136, -455	1, 14, 24, 13, 91, 276, 364
-1, -30, -26, -17, -435, -325, -4060	1, 10, 25, 13, 45, 300, 120
-1, -45, -25, -17, -990, -300, -14190	1, 11, 33, 15, 55, 528, 165
....	....

After solving system (3.3) for the training sequence, which consists of 15 representatives of class I and 22 of class II, and then after computing vector Y according to formula (3.2), we will obtain: $y_0=2483$; $\Delta_1 x=-0.05$; $\Delta_1^2 x=-2.5$; $\Delta_1^3 x=-16.5$; $\Delta_2 x=-14$; $\Delta_2^2 x=22$; $\Delta_3 x=-1$.

Thus, the code table will take the following form (fragments of the tables are brought). Here are selections $x_{10}=6833$; $x_{20}=4326$; $x_{30}=-24$. Values are rounded off to whole numbers.

(a) Номер	(b) Значение параметра	(c) Код	(a) Номер	(b) Значение параметра	(c) Код	(a) Номер	(b) Значение параметра	(c) Код
(d) Коды ρ_k			(d) Коды ПС			(e) Коды кавернограммы		
1	1	+6833	1	—20	—4340	1	25	—25
...	...	....	...	...	....	...	...	...
10	10	+4740	23	2	+918	13	37	—37
11	11	+3973	24	3	+1410	14	38	—38
12	12	+3038	25	4	+1924	15	39	—39
13	13	+1919	26	5	+2460	16	40	—40
14	14	+599	...	...	....	17	41	—41
15	15	—937				...	...	...
...	...	....						
30	30	—61246						
...	...	....						

Key: ***(a). Number. (b). Value of the parameter. (c). Code. (d). Codes ....***
(e). Caliper Log Codes.

For determination, does relate object 13, 25, 15 to the frozen or thawed rocks, it is necessary to take the codes from the tables, which correspond to 13 for ρ_k, 25 for PS and 15 for the caliper log and add them together. If the sum is negative, then this layer relates to frozen, if it is positive - to thawed. In particular, the object indicated relates to thawed rocks.

5. Recognition of objects. It is possible to use the code table $\{x_{ij_i}\}$ for the classification of objects, not included in the training sequence. For this it is sufficient to determine the numbers of the values, which each parameter accepts, according to the table to find the codes corresponding to them and to add. If the sum obtained is negative, then object relates to the 1st class, if positive - to the 2nd.

Tables for predicting the oil content of layers according to geophysical data measurements were obtained in the Perm' university computer center by the method described above, also research on the application of these methods is carried out for predicting the properties of double metallic systems.

The following properties were forecast: the presence in a binary system of a chemical compound of a given composition and type, diagrams of the condition in the vicinity of a given compound.

Page 147

The electron structure of the atoms was examined as a description of the objects, which compose system. The dividing plane was conducted in expanded parameter spaces, where the pair combinations of the similar parameters of the system (a system consists of two components) were used. During the construction of the code table the training sequence was completely divided, but the systems used for the examination, which were not included in the training sequence, were classified correctly in 70% of the cases. The basic results of these research are given in works [5-7].

REFERENCES

1. Yu. V. Devingtal'. On optimum coding of objects with classification by methods of pattern recognition. Izv. AS USSR, Technical cybernetics, 1968, No. 1.

2. M. M. Bongard, Sh. A. Guberman. Program of the recognition of oil-bearing layers "Kora-3". In collection: "Problems of the expansion of the possibilities of automata", iss. 5, materials seminars, Institute of automation and telemechanics, 1965.

3. Yu. V. Devingtal'. The solution algorithm of system of inequalities, to which the

problem of the classification of objects is reduced. In collection: "Algorithms and programs for computer Aragats" Scientific notes PGU, No. 220, Perm', 1970.

4. S. I. Zukhovitskiy, L. I. Avdeyeva. Linear and convex programming. "Science", 1964.

5. Ye. M. Savitskiy, Yu. V. Devingtal', V. B. Grabulya. On recognition of double diagrams of state of metallic systems with computer aid. DAN SSSR, 178, No. 1, 1968.

6. Ye. M. Savitskiy, Yu. V. Devingtal', V. B. Gribulya. Prediction of metallic compounds A_3B with computer aid, DAN SSSR, 183, No. 5, 1968.

7. Ye. M. Savitskiy, Yu. V. Devingtal', V. B. Gribulya. Recognition of the type of formation reactions and evaluation of the interval of the homogeneity of metallic phases with computer aid. DAN SSSR, 1985, No. 3, 1968.

Submitted 19 May 1969

DISTRIBUTION LIST

DISTRIBUTION DIRECT TO RECIPIENT

ORGANIZATION	MICROFICHE
B085 DIA/RTS-2FI	1
C509 BALLOC509 BALLISTIC RES LAB	1
C510 R&T LABS/AVEADCOM	1
C513 ARRADCOM	1
C535 AVRADCOM/TSARCOM	1
C539 TRASANA	1
Q592 FSTC	4
Q619 MSIC REDSTONE	1
Q008 NTIC	1
Q043 AFMIC-IS	1
E051 HQ USAF/INET	1
E404 AEDC/DOF	1
E408 AFWL	1
E410 ASDTC/IN	1
E411 ASD/FTD/TTIA	1
E429 SD/IND	1
P005 DOE/ISA/DDI	1
P050 CIA/OCR/ADD/SD	2
1051 AFIT/LDE	1
P090 NSA/CDB	1
2206 FSL	1

Microfiche Nbr: FTD93C000345L
FTD-ID(RS)T-0338-93